Analysis of liability constraints

Cuijpers, C.M.K.C.; Schellekens, M.H.M.; Leenes, R.E.

Publication date:
2011

Document Version
Publisher's PDF, also known as Version of record

VERSATILE INFORMATION TOOLKIT FOR END-USERS ORIENTED OPEN-SOURCES EXPLOITATION

Funding Scheme: CP - Collaborative Project

Call (part) identifier: FP7 - SEC - 2009 - 1

H 3.2.4 – ANALYSIS OF LIABILITY CONSTRAINTS

Project Number: FP7 - SEC - GA - 2009 - 242352

Project Title: Versatile Information Toolkit for End-Users Oriented Open-Sources Exploitation

Documento ID: VIR-3009-WP3-007-V01-TW-CO **Revision:** Draft

Date of Minutes: 30/09/2011

Project co-funded by the European Commission within the Seventh Framework Programme (2007-2013)		
Dissemination Level		
PU	Public	
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	CO

Revision	**Edition Date**	**Author**	**Modified Sections / Pages**	**Comments**
0.1	30.09.2011	Colette Cuijpers	First draft	
0.2	07.10.2011	Maurice Schellekens	First draft	
0.3	20.10.2011	Colette Cuijpers	First draft	
0.4	26.10.2011	Ronald Leenes	First draft	
0.5	31.10.2011	Colette Cuijpers	Final draft	

Table of Contents

This project has been carried out under a contract awarded by the European Union

1. Introduction

1.1 Background

This heartbeat gives an insight into how liability risks relating to possible legal and ethical infringements caused by (the use of) the Virtuoso framework can be reduced. Infringement of ethical and legal constraints may give rise to liability issues on several different levels. Designers, builders, the commissioner of the tool as well as its end users all perform actions which can lead to damages for which liability might exist.

> This document will not address liability issues within the Virtuoso consortium. As Virtuoso is a research project, aimed at building a prototype and not an actual product phase application, flaws in the work carried out in the different work packages which might hamper a proper functioning of the desired outcome of Virtuoso, will be regarded as part of ongoing research and not as actions that lead to product liability. Even though not performing according to the Consortium Agreement might lead to liability for the defaulting party, no liability issues outside the scope of the research consortium are to be expected.

This heartbeat focuses on the liability issues that arise in relation to the Virtuoso framework as such, its components, the prototype and real life application in relation to society as a whole. In this respect questions are addressed that relate to:

1. Does the mere coming into existence of a Virtuoso framework raise liability concerns and if so, (how) can Virtuoso deal with these issues? In this respect the platform, the components and the demonstrator are assessed separately.

2. (How) can Virtuoso cope with liability issues that arise in view of end-users application of the Virtuoso framework?

1.2 Purpose and Scope

This heartbeat aims to clarify how the Virtuoso Consortium or Virtusos partners can deal with liability issues that might arise in respect of the different layers of the Virtuoso framework. The scope of this heartbeat is determined by two important restrictions that form part of the overall Virtuoso research project:

1. End users are restricted to public authorities in the field of intelligence.
Legal obligations and rights can differ heavily in relation to public parties on the one hand and private parties on the other. Possibilities for private parties to legitimately use Virtuosos functionalities will probably be far less than for public authorities. For instance, the processing of sensitive data (as defined in Article 16 of the 95/46/EC Data Protection Directive) by private parties is in essence prohibited, but Member States are left the option to create exceptions in view of ‘substantial public interest’ (reasoning from the situation that Virtuoso technology will be unable to avoid the processing of sensitive information when mining open sources). In view of intellectual property, for private parties using Virtuoso commercially, no exceptions exist that justify the reproduction of copyrighted materials. Yet, by using Virtuoso, the reproduction of copyrighted materials is inevitable.

This project has been carried out under a contract awarded by the European Union

2. The use of Virtuoso is restricted to open source information, in a strict sense. This is defined as follows. Information is unclassified and derived from overt, non-clandestine and non-secret sources and is publicly available to anyone without breaking any kind of technical or organisational protection measure. It can be lawfully obtained by request, purchase, or observation (Best and Cumming 2007: 8-10). In this respect it is interesting that Best and Cumming claim that the acquisition of such information *'must conform with extant copyright requirements'*. In this heartbeat several guidelines are given how Virtuoso can cope with copyright issues.

1.3 Document Structure

As described in Heartbeat 3.2.1, and further developed in the WP3 wiki on www.virtuoso.eu, for the purpose of the legal and ethical analysis we distinguish between several layers to which the legal and ethical analysis applies differently. The legal and ethical analysis concerns four different layers of the Virtuoso framework (the term used to refer to the complete set of Virtuoso layers).

1. Platform. The platform consists of a relatively empty shell. It is composed of four different frameworks (processing, evaluation, decision support and presentation) and it includes factories to ease integration of new components. The components are not part of the platform. One of the important questions to be addressed is, to what extent, the platform can be viewed in complete disconnection from the components.

2. Components. Pieces of software that will be integrated to achieve desired functionality of the platform. Which components are necessary to achieve certain functionality depends on the requirements of the end product. Components can be existing ones that are being used in the course of the project as well as newly developed components specifically designed for Virtuoso.

3. Prototype. The prototype consists of an instance of the platform with several different components. It is a working system that, when actual information is being processed with it, demonstrates the full capabilities of the Virtuoso vision. The prototype will be developed as a laboratory setting. The data being used may be real open source data (or manufactured data), but the end users will be simulated rather than the real (intended) end users.

4. Application. This concerns the prototype outside a laboratory setting. In this scenario the end users will become responsible for the use of the Virtuoso platform and its components. However, it is a question open for debate whether in such a scenario Virtuoso as a developer (and possibly as a provider of (maintenance services) may also carry some responsibility.[1] Is it purely up to the end users to comply with the existing legal and ethical framework?

The legal and ethical research regarding Virtuoso is furthermore clustered in 6 domains: privacy and data protection, (other) human rights, ethics, intellectual property, investigative authorities and liability. Liability is a consequence of actions (usually infringements of rights) in the five other domains.
Outlining liabilities on a European level is impossible, even when restricted to the sphere of investigative authorities, because international and/or European harmonisation is lacking. Instead, regulation within individual Member States has to be observed. All Member States have authorities in place granting legal room for intelligence and law enforcement to monitor and analyse open source information.

[1] A car manufacturer, for instance is liable for damages caused by its drivers when caused by certain malfunctions of the car. Without the action of the driver, the damages would (possibly) not have occured, yet the driver might have been able to prevent the damages if the car was not faulty.

The main structure of this document traverses the 4 identified layers of the Virtuoso framework. In subsections, the liability issues are described relating to human rights and ethics, data protection and intellectual property rights. Privacy is grouped under human rights instead of being grouped with data protection because the first analysis of the platform did not reveal issues of data protection, but did reveal privacy issues. Because other heartbeats already shed light on the legal ethical domains from which liability risks emerge, and in view of the desire of the technical partners to have concise 'checklists' instead of extended research papers, this heartbeat provides a list of questions that need to be answered to avoid liability risks.

The heartbeat will conclude with some remarks that, together with questions provided for by the technical partners, will be used as input for a WP3 privacy & data protection seminar that will take place in Brussels on 22 November 2011.

1.4 Applicable and Reference Documents

As already described, this heartbeat follows the typology/taxonomy that is created in the form of a wiki (available at www.virtuoso.eu). This work in progress, to which all partners in the project are invited to contribute to, helps to further the understanding of the complex issues, situations and concepts that are being dealt with in the Virtuoso project.

This heartbeat builds upon the legal and ethical quick scan performed in Heartbeat 3.2.1. Together with Heartbeats 3.2.2, 3.2.3, 3.2.5, 3.2.6 and Deliverable 3.1, this heartbeat constitutes one of the building blocks for Deliverable 3.2, the analysis of the legal and ethical framework in open source intelligence.

2. Liability for the platform

It is a misconception that, because the platform in itself 'does nothing', it is not subject to any legal and/or ethical constraints. To illustrate this, reference can be made to providers of collections of links to music or video downloads, who are held liable by some courts for copyright infringement even though they do not provide actual content for download themselves.
In this section several questions will be presented that need to be answered properly before the conclusion can be drawn that 'there are no liability issues with the Virtuoso platform'.

Question 1: *To what extent can the platform be viewed as a distinct entity disconnected from the components.*

2.1 Human rights and ethics

Even though this section is about the Virtuoso platform, it is necessary to stress that the questions that will be dealt with in this subsection concern what we call the Virtuoso framework as comprised by the four Virtuoso layers in a 'non-operational-state'. The questions do not relate to a functional instance of such a framework, but rather concerns the abstract level of 'mere existence' of such a framework. From the viewpoint of human rights and ethics it is important to address the effect of the existence of a Virtuoso framework within society (when able to function in practice as intended). Knowing that such a framework is in place can infringe upon people's (sense of) privacy, freedom and security. If I know (and the I in this sentence can also relate to someone seeking asylum or a visa) that all the data that I myself, or other people, reveal about me in open sources is monitored by intelligence authorities throughout Europe, will I adapt my behaviour or feel hampered in e.g. my freedom of expression, or do I feel my privacy being invaded? This relates to Benthams panopticon, in which the mere possibility of being watched affects behaviour. Foucault describes this as: *"He who is subjected to a field of visibility, and who knows it, assumes responsibility for the constraints of power; he makes them play spontaneously upon himself; he inscribes in himself the power relation in which he simultaneously plays both roles; he becomes the principle of his own subjection"* (Foucault 1977, discipline 202-203).

From the analysis of ethics and human rights it becomes clear that liability issues not only emerge because private or public parties infringe upon ethical or legal rights by action, but also by non-action. Member States can also be liable if they do not take the necessary measures to safeguard a right or, more precisely to adopt reasonable and suitable measures to protect the rights of the individual (Akandji-Kombe: 5).

In order to conform to ethical and human rights standards it is important to assess the overall desirability and legality of a framework like Virtuoso. In this respect criteria like 'proportionality', 'subsidiarity' and 'necessity' flow from the ethical and human rights requirements. If infringements of human rights are expected, which is likely regarding privacy, it will boil down to the assessment whether the infringement is justified. From this perspective it is necessary to show that Virtuoso actually contributes to the aim for which it is designed. Otherwise, it is hard to stand the test of 'necessity', 'subsidiarity' and 'proportionality'. In view of the issue of liability, this entails that within the Virtuosos research project careful thought must be given on how to answer questions like:

Question 2: *Is a framework like Virtuoso necessary in a democratic society, art. 8 para 2 ECHR?*

This project has been carried out under a contract awarded by the European Union

Phrased differently: What is the justification for developing such a framework especially as the development is funded with community money?

Question 3:
How is this need substantiated? (e.g., scientific research, empirical data etc.)

Question 4:
Is a framework like Virtuoso proportional in view of its aim?

Question 5:
Is it capable of achieving its aim? How can the claim that it will contribute to 'border control' be substantiated?

Question 6:
Are there less invasive ways to achieve the perceived aim?

Question 7:
What measures have been taken to make Virtuoso (better) conform to proportionality and subsidiarity standards?

In view of the answers to the previous questions, it is important that Virtuoso is restricted in two senses. First, the end users should be limited to public authorities in the field of intelligence (even more strict: border control) in order to stay within the boundaries imposed by for instance, art 8 para 2 EHCR. Second, data gathering and analysis should be restricted to data in open sources because this limits the (privacy) infringements. These two restrictions contribute to a positive analysis (in a sense: no liability caused by surpassing ethical or human rights constraints) but also raise new issues in relation to how these restrictions are guaranteed.

Question 8:
What guarantees are in place to make sure the framework is only used by its intended end-users?

Question 9:
What guarantees are in place that Virtuoso will restrict itself to mining open sources?
More specific: *How is it guaranteed that the policies outlined in robot.txt files will be honoured and how is it guaranteed that secured information (e.g., password protected files) is ignored by Virtuoso?*

More in general:

Question 10:
How is (mis)use of Virtuoso (by private parties) secured? (As Virtuoso is meant to be open source software, see question 15).

Whether or not a lack of such guarantees will contribute to liability issues depends on the risks that are associated with 'other end-users making use of Virtuoso' and 'other than open source data being used by intended end-users' or the combination 'other end-users making use of other than open source information'.

Question 11:
What are the risks if Virtuoso is used by others than the intended end-users? (see heartbeat 3.2.6 risk analysis).

Question 12:
What are the risks of Virtuoso extending beyond open source information? (see heartbeat 3.2.6 risk analysis).

Question 13:
What are the risks of others than the intended end-users using Virtuoso beyond open source information? (see heartbeat 3.2.6 risk analysis).

If illegal use of the platform or certain components is to be expected, the design of the framework must (at least try to) prevent this illegal use, or restrict it as much as possible (preferably by making it impossible or extremely hard, e.g. by technical safeguards).

It is important to stress that it is questionable whether liability for illegal use of the platform and its components can completely be shifted to the Virtuoso end-users. First of all, in analogy to e.g. file sharing software and services, a product and/or service, must have at least substantial non infringing uses.[2] Moreover, in analogy with e.g. online intermediaries, if knowledge exists of illegal use of one's products or services by third parties, and, even though one is capable of intervening, no action is taken to prevent or stop illegal behaviour, this might lead to liability of the behaviour of the intermediaries.[3]

On the basis of purpose and exceptions (in the sphere of public interest, national security etc.) we presume Virtuoso can be legally used by end-users (law enforcement officials) if adhered to the ethical and legal requirements identified in WP3. As these exceptions mostly stem from national legislation, a separate analysis for each and every Member State is necessary to determine what is and what is not allowed under what conditions and safeguards. This goes beyond the scope of the research. We will therefore adopt the (reasonable) general assumption that national intelligence authorities will have the authority to investigate open sources. In view of private parties the legitimacy of mining open sources is harder to substantiate because they can not base their use on the contribution to public interest when using Virtuoso. Therefore, the question whether there is fair and legal use of Virtuoso by private parties is more difficult to answer. This understanding leads to the following questions:

Question 14:
(How) can Virtuoso be used by private parties without infringing copyright and or privacy legislation?
What legitimate and fair uses of Virtuoso by private parties are envisionable?

Question 15:
If Virtuoso has no substantial non-infringing uses for private parties, is release under open source software allowed/the best way to proceed?

[2] Betamax-case: Sony Corp. of America v. Universal City Studios, Inc., 464 U.S. 417 (1984). Even though this concerns US case law, it is refered to in view of EU law as well.

[3] See art. 14 of the E-commerce Directive (2000/31/EC). Examples of cases where it was judged that the intermediary did not correctly intervene: Cour d'Appel de Paris 14ème Chambre, Section A Arrêt du 12 Décembre 2007, Google Inc v Benetton, Bencom, available at http://legalis.net/spip.php?page=jurisprudence-decision&id_article=2116 and Tribunal de grande instance de Paris Ordonnance de Référé 29 Mai 2007, Benetton, Bencom v Google Inc, Google France (2007), available at http://legalis.net/spip.php?page=jurisprudence-decision&id_article=2120 (accessed 28 July 2011).

2.2 Data protection

If the bare platform does not process personal data, liability for infringement of data protection regulations is not to be expected at this level.

2.3 Intellectual property rights

If the platform does not as such 'reproduce' or 'make available to the public' copyright protected materials, no liability claims are to be expected.
Intellectual property is not only protected by copyright law, but can also be protected by technical protection measures (TPMs), such as anti copy protection. These TPMs in turn are legally protected in a number of copyright directives which prohibit the circumvention of technical copy- or access protection. This legal protection does not just target the act of circumvention itself, but also putting into circulation and the possession for commercial purposes of the means for circumvention (for precise definitions see art 6.2 Directive 2001/29/EC and art. 7.1(c) Directive 2009/24/EC). If the platform includes means to circumvent technical protection, the mere possession for commercial purposes (such as offering for sale or rental) of the platform may fall afoul of legislation implementing the abovementioned directives.

Question 16:
Does the platform provide for means to circumvent effective technological protection measures as defined in art. 6.3 of Directive 2001/29/EC (Copyright in the Information Society) or a means for removal or circumvention as meant in art 7.1 (c) Directive 2009/24/EC (Software Protection)?

The intellectual property issues relating to ownership of the platform as such are secured in the contracts underlying the Virtuoso project and are not part of the legal and ethical research of WP3.

3. Liability for components

From the perspective of human rights and ethics a similar line of reasoning and questioning is applicable to components as it is to the platform (framework) as such. Therefore no separate analysis will be given of the components. They are considered to be 'just tools' that do not process data as such. If testing the functionalities of components concerns the actual processing of data, this will be dealt with in section 4 regarding liability issues surrounding the prototype. Consequently, we will not provide a separate analysis of liability issues relating to data protection and intellectual property issues because we assume that the components as such do not infringe upon these rights because no actual data are being processed. Where data is processed, we will address the issues in section 4 which deals with the prototype.

The intellectual property issues relating to ownership of the components as such are secured in the contracts underlying the Virtuoso project and are not part of the legal and ethical research of WP3.

4. Liability for the prototype

4.1 Human rights and ethics

In relation to the prototype, the perspective of human rights and ethics is probably less important than it is in relation to the Virtuoso framework discussed above. The prototype as such has as its sole purpose to

demonstrate that the Virtuoso framework is able to perform the functions ascribed to it. The need to demonstrate functionality is an essential part of validating research results and as such not very controversial. However, just as is the case with the framework as a whole, safeguards need to be in place to limit the functionality and use of the prototype to its purpose; demonstration of the functionality of the Virtuoso platform and its components. There are several restrictions to the prototype that will prevent triggering the question whether citizens will feel impaired in their freedom when they know of a Virtuoso prototype being used to test the functionality of the Virtuoso platform and its components. First of all the prototype is used in relation to three pre-defined scenario's, meaning there will not be undirected searches of open sources. Second, the duration of mining the open sources is limited in time. The use of the prototype will terminate once the demonstration has proved Virtuoso to function. Third, and probably most important, there will be no actual take up on the basis of the outcomes. So no decisions will be based on the demonstration results. Therefore, looking at heartbeat 3.2.2 regarding human rights and ethical constraints, the human right to be most relevant in view of the current analysis is the right to privacy, and more specific the right to informational privacy. The implications of data protection legislation on the prototype will be discussed in the next section. From a viewpoint of the more general right to privacy, the test of article 8 ECHR comes into play. Does the prototype infringe upon the right to privacy (art 8 para 1 EHCR)? If actual data are being used, and an actual analysis is provided for in the demonstration, this will probably involve actual people that might feel their privacy is invaded, even though the purpose is only to demonstrate the functionality of Virtuoso. The second step in the article 8 test concerns the justification of the infringement (para 2). As a legal basis, reference can be made to the *protection of the rights and freedoms of others,* one of the interests mentioned in art. 8 (2). However, the interesting questions stem from the principles of proportionality and subsidiarity. The specific questions that arise in relation to data protection are dealt with in the next section, leaving open the more general question:

Question 17
What measures have been taken to limit the legal and ethical implications of the prototype?

Question 18
What measures have been taken to limit the application of the prototype to its intended use: purely demonstrate Virtuoso functionality?

An essential difference in relation to the scenario in which Virtuoso is actually used by end-users, is the fact that this research takes as a starting point end-users that fall within the broad notion of 'public authorities in the field of 'intelligence'. This limitation is grounded in the fact that national legislation provides for these types of end-users exceptions to several existing legal constraints, in the field of data protection as well as in the field of intellectual property. Regarding the prototype it is questionable that its users can be brought under the legal exceptions that are put in place by national authorities in view of intelligence authorities because the users of the prototype will be partners within the consortium which are mainly private (commercial) parties as well as semi-private institutions such as Universities. This means that other legal regimes, and with that other (if any) exceptions, are applicable to these users. In the next sections it will be discussed how this works out in view of data protection and intellectual property.

4.2 Data protection

In developing the prototype, probably the first issue that arises in relation to data mining of open sources is the possible conflict with data protection legislation. Article 2(a) of Directive 95/46/EC states; *"'personal data' shall mean any information relating to an identified or identifiable natural person ('data subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic,*

cultural or social identity". It is unlikely that technology is able to determine for each and every piece of information in open sources whether or not it legally qualifies as personal data. Especially as combinations of trivial data may lead to identifiability, and thus qualify as personal data. Moreover, from an Article 29 Working Group Opinion it can be deducted that data must be considered to be personal data when identifiability is uncertain, however possible. A similar reasoning is put in place regarding WiFi data: *"Under these circumstances and taking into account that it is unlikely that the data controller is able to distinguish between those cases where the owner of the WiFi access point is identifiable and those that he/she is not, the data controller should treat all data about WiFi routers as personal data"*.[4] So, in mining open sources personal data are being processed and thus the regime regarding processing of personal data is applicable, unless exceptions apply.

Problematic in view of mining open sources is the inherent tension between limiting the secondary use of data and mining data that was collected for purposes other than data mining. This leads to the obligation for those conducting data mining activities to examine whether the mining of data is consistent with the purposes for which those data were originally collected. In view of sources like the Internet, this definitely is an unrealistic task.

Here we touch upon a more fundamental issue, and what we could call a grey area, as a friction can be identified between what is actually happening in practice and what is allowed on the basis of strict interpretation of data protection legislation. If we take a look at the Internet, we can identify several websites and tools that are in the business of mining open sources. To find other people websites like www.pipl.com, www.wieowie.nl and the software package Maltego (http://www.paterva.com/web5/) all provide for possibilities to search open sources, mostly to find specific individuals and information relating to them. Virtuoso will take this process several steps further, but in essence it concerns the same type of exercise. However, the legal basis for these type of searches, when they can not be placed under the 'purely household exception', is questionable. Private parties can only reside to the legal ground: *'processing is necessary for the purposes of the legitimate interests pursued by the controller or by the third party or parties to whom the data are disclosed, except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject which require protection under Article 1 (1)'*. Maybe this ground can be upheld in view of the mere mining of the open source, but when the purpose becomes intertwined with intelligence and forms part of the decision making processes in relation to border control issues, it is definitely open for discussion whether privacy interests should prevail or not.

Public authorities can base open source searches on the legal ground 'public interest': *'processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller or in a third party to whom the data are disclosed'*.

With a view to processing sensitive data, a qualification that is hard to detect by data mining technology, the legitimacy of data processing might even be more problematic. The processing of sensitive data is prohibited unless an exception is in place. Again for public authorities national legal regimes will most certainly provide for proper exceptions as Art. 8 (4) of Directive 95/46/EC explicitly leaves this option open: *'Subject to the provision of suitable safeguards, Member States may, for reasons of substantial public interest, lay down exemptions in addition to those laid down in paragraph 2 either by national law or by decision of the supervisory authority'* and Art. 8 (5): *'Processing of data relating to offences, criminal convictions or security measures may be carried out only under the control of official authority, or if suitable specific safeguards are provided under national law, subject to derogations which may be granted*

[4] Art. 29 Working Party Opinion 13/2011 on Geolocation services on smart mobile devices (WP 185), p. 11, Available from: http://ec.europa.eu/justice/policies/privacy/workinggroup/wpdocs/2011_en.htm

by the Member State under national provisions providing suitable specific safeguards. However, a complete register of criminal convictions may be kept only under the control of official authority.
Member States may provide that data relating to administrative sanctions or judgements in civil cases shall also be processed under the control of official authority'.

In view of private parties the only grounds that might offer relief are Art. 8 (a) and (e):

'the data subject has given his explicit consent to the processing of those data, except where the laws of the Member State provide that the prohibition referred to in paragraph 1 may not be lifted by the data subject's giving his consent.'

'the processing relates to data which are manifestly made public by the data subject or is necessary for the establishment, exercise or defence of legal claims'.

For both exceptions it is problematic that data in open sources is not necessarily made available by the data subject. A lot of personal data in open sources will be posted by third parties. Moreover and as already referred to, are data mining activities within Virtuoso consistent with the purposes for which those data were originally collected. Principles like purpose limitation and data minimisation constrain further use of data. In relation to consent it can also be mentioned that even when a person discloses information about himself online, that this does not entail consent for everyone to use this information in any way desired. Consent can only be given freely and for well defined specific purposes.

From the foregoing it becomes clear that data protection legislation might be more problematic in view of the use of the prototype, then it is in respect of actual end use by public authorities. In view of the strict interpretation given to the household exception (This Directive shall not apply to the processing of personal data by a natural person in the course of a purely personal or household activity, Art. 3 (2) Directive 95/46/EC) this exception does not give legal room for Virtuoso.[5] While the prohibition to process sensitive data is hard to overcome by any other exception. Even the general exception clause (Art. 13 of Directive 95/46/EC) does not refer to the processing of personal data (which is regulated in art. 8):

"1. Member States may adopt legislative measures to restrict the scope of the obligations and rights provided for in Articles 6 (1), 10, 11 (1), 12 and 21 when such a restriction constitutes a necessary measures to safeguard:
(a) national security;
(b) defence;
(c) public security;
(d) the prevention, investigation, detection and prosecution of criminal offences, or of breaches of ethics for regulated professions;
(e) an important economic or financial interest of a Member State or of the European Union, including monetary, budgetary and taxation matters;
(f) a monitoring, inspection or regulatory function connected, even occasionally, with the exercise of official authority in cases referred to in (c), (d) and (e);
(g) the protection of the data subject or of the rights and freedoms of others.
2. Subject to adequate legal safeguards, in particular that the data are not used for taking measures or decisions regarding any particular individual, Member States may, where there is clearly no risk of breaching the privacy of the data subject, restrict by a legislative measure the rights provided for in Article 12 when data are processed solely for purposes of scientific research or are kept in personal form for a period which does not exceed the period necessary for the sole purpose of creating statistics."

[5] JUDGMENT OF THE COURT, 6 November 2003, in Case C-101/01 (Bodil Lindqvist). OJ C 118 of 21.4.2001.

In a large scale EU research regarding new challenges to data protection, the issue of open source information is discussed as an example of how the legal landscape regarding data protection framework can come into conflicto with mere and existing practice (Korff 2010). A line of reasoning that can be identified concerns the balancing of the right to freedom of speech against the right to privacy.

In view of the fact that this concerns a grey legal area, an area for which no black and White answers can be given at this momento in time, WP3 will invest in a scientific article in which this dicussion will be fully addressed. In Deliverable 3.2, of which this heartbeat is one of the building blocks, preliminary findings regarding the grey legal area will be provided for. The article is to be expected March 2012.

For now it is enough to point to the difficulties that emerge from data protection legislation in view of processing sensitive personal information in mining open sources, especialy when it concerns this practice in relation to the prototype. One exercise that is essential in this respect is specifying the purposes for data processing in the realm of the prototype.

Question 19
What exactly is the purpose of the data processing taking place in view of the prototype?

A general description is not sufficient in this respect. Moreover, the purpose of each and every processing of personal data taking place in the course of developing Virtuoso must be specified, explicit and legitimate and must be determined prior to the actual collection of the data.

Question 20
Has the purpose been specified for each and every processing of personal data that takes place in the development of Virtuoso?

4.3 Intellectual property rights

Experimental use of the prototype is basically governed by the same IP laws that govern the use of the finished product. What has been described in section 3.2.3 therefore applies here as well. However, experimental use of the prototype may benefit from exceptions that allow for use of copyrighted works for scientific purposes. The 'scientific purpose' exceptions are bound to limits and conditions. Do these limits and conditions prevent application of the exceptions to experimental use of the prototype? We start our analysis with article 5 section 3 of Directive 2001/29/EC (Copyright in the information Society[6]) :

> 3. Member States may provide for exceptions or limitations to the rights provided for in Articles 2 and 3 in the following cases:
>
> *"use for the sole purpose of illustration for teaching or scientific research, as long as the source, including the author's name, is indicated, unless this turns out to be impossible and to the extent justified by the non-commercial purpose to be achieved".*

[6] Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society, OJ L 167, 22.6.2001, p. 10–19.

The directive basically allows for a broad exception for scientific research. It does however neither contain an obligation for Member States to implement the provision in national law, nor an obligation to create an exception that is as wide as the exception proposed in art. 5.3 (a) Dir 2001/29/EC. We have examined a number of selected national provisions – implementing art 5.3 (a) Dir 2001/29/EC – in order to assess their application to experimental use of the prototype. National provisions in the following states have been examined: France, Germany, Italy, the Netherlands, Spain and the UK. An English translation of the provisions can be found in the annex. The selected national provisions – partially – implementing the exception as described in art. 5.3 (a) Directive 2001/29/EC show shortcomings that make them less or even unsuitable for use in relation to the Virtuoso prototype. Shortcomings include:

- The exception does not relate to the reproduction right: e.g. Germany, and the Netherlands. In Italy it does not relate to 'permanent reproduction'.
- The exception only relates to certain forms of 'making public', such as recitation, public performance or presentation. This suggests that the exception is only applicable where the public is gathered at the place of recitation etc.: e.g. the Netherlands.
- The exception only relates to education, not to research: Germany.
- Application of the exception requires a remuneration to be paid to the rightsholder (art 52 German Copyright Act).
- The exception is limited to certain types of works: e.g. databases (art. 64sexies Italian Copyright Act).
- Some countries do not have exceptions relating to research, e.g. France and Spain.

The implementing provisions in the examined Member States show significant differences. There is no true harmonisation of copyright exceptions for research purposes. Some countries have no relevant exception at all. The UK has a wide exception, allowing for all fair dealings with works for non-commercial research. Other countries have exceptions predicated on conditions that make the exceptions of little use to the prototype.The research exemption has serious drawbacks. Other means to provide a legitimisation for use of protected works have drawbacks also. It is unlikely that an implied license extends over re-use of copyrighted materials. Even small usage of materials such as thumbnails was found not to be covered by an implicit license.[7] In the Google-Copiepresse case, it was explicitly decided that an implicit license could not be derived from the non-usage of metetags or the robot.txt protocol by the rightsholders.[8] In a recent decision (19 October 2011), the German Bundesgerichtshof reputedly widened the possiblities for search engines to use thumbnails relying on an implicit license.[9] An implicit license to use a thumbnail of a picture in a search engine was found where the rightsholder had given anybody permission to place the picture online, where the picture was available without protection measures being in place and it did not matter that the search engine's thumbnail linked to a website that had no permission to use the picture. The text of the decision was not available at the time of writing this text. In conclusión, it is incertain that an implied license can be relied upon for processing or translating works or for making them available to the public. Obtaining explicit licenses would provide a legitimisation, but may be burdensome to obtain. The above results in the following liability risks:

- Running the prototype or conducting activities related to running the prototype in countries where no or a narrow research exception is in place.

[7] Hamburg Regional Court 308 O 42/06 (Thomas Horn-Google) and 308 O 248/07 (Michael Bernhard-Google).

[8] Brussels Court of Appeal 5 May 2011 Google-Copiepresse at no. 50 and 51, available at: http://www.ie-forum.nl/backoffice/uploads/file/IEF%20Hof%20van%20Beroep%20Brussel%205%20mei%202011,%20R_nr_%202011_2999,%20817%20(Google%20INC_%20tegen%20Copiepresse,%20Societe%20de%20Droit%20d'auteur%20des%20journalistes%20(SAJ)%20en%20Assucopie).pdf (visited 28 October 2011).

[9] See http://juris.bundesgerichtshof.de/cgi-bin/rechtsprechung/document.py?Gericht=bgh&Art=pm&pm_nummer=0165/11 .

- Transgressing the boundaries of an implied license, especially if an implied license is relied upon to provide cover for activities such as processing or translating works or communicating them to the public or where it concerns works placed online illegitimately.
- Running the prototype or conducting activities related to running the prototype without adequate licenses where other justifications for the uses of works are lacking or uncertain.

5. Liability for actual application

The reasoning regarding liability for the use of the prototype is to a large extend applicable to the actual end use of Virtuoso. Depending the characteristics of the actual use (purpose, by whom, how long, etc.) the impact of actual use is most certainly much more invasive then testing the prototype. However, the legal and ethical tests regarding legitimacy of the use in essence remain the same, even though the outcomes may differ, e.g. the test whether use is in conformity with principles of proportionality and subsidiarity. Even though it is shown in this heartbeat that it is not possible to entirely shift all liability issues to the end-users, if adhered to the requirements, guarantees and obligations described in this heartbeat, liability for actual use will mainly reside with the end-user. In view of this, we will not further elaborate on liability for actual use.

6. Conclusion

Liability risks emerge when legal requirements are not correctly adhered to or when insufficient measures have been taken to prevent (the use of) Virtuoso causing damage. This heartbeat provides a list of questions which, if properly answered and complied with, can help reduce liability issues.

Main issues to be dealt with:

1. Substantiate need, proportionality and subsidiarity of Virtuoso
2. Create mechanisms to comply with legal rules and regulations
   - technical (e.g. anonymisation tools, encryption tools)
   - legal/organisational (e.g. design processes for correction of data, deletion of data etc.)
3. Create mechanisms to lessen risks of misuse
   - technical (authentication mechanisms, data security etc.)
   - legal/organisational (information for end users, contracts, etc.)
4. Further research regarding the grey area of practice versus strict application of the legal framework

References

Akandji-Kombe, Jean-François (2007) *Positive obligations under the European Convention on Human Rights. A guide to the implementation of the European Convention on Human Rights*, Human rights handbooks, No. 7, http://www.echr.coe.int/NR/rdonlyres/1B521F61-A636-43F5-AD56-5F26D46A4F55/0/DG2ENHRHAND072007.pdf

Bentham, J. Panopticon (Preface) (1995) in Miran Bozovic (ed.), *The Panopticon Writings*, London: Verso, 1995, 29-95.

Best, Richard A. and Cumming, Alfred (2007) *Open Source Intelligence (OSINT): Issues for Congress*, Order Code RL34270, December 5, 2007, http://www.fas.org/sgp/crs/intel/RL34270.pdf.

Foucault, M. (1977) *Discipline and Punish: The Birth of the Prison*. Trans. Alan Sheridan. New York: Pantheon.

Korff, Douwe, New Challenges to Data Protection Study - Working Paper No. 2: Data Protection Laws in the EU: The Difficulties in Meeting the Challenges Posed by Global Social and Technical Developments (January 15, 2010). European Commission DG Justice, Freedom and Security Report. Available at SSRN: http://ssrn.com/abstract=1638949

Korff, Douwe and Brown, Ian, New Challenges to Data Protection - Final Report (January 20, 2010). European Commission DG Justice, Freedom and Security Report. Available at SSRN: http://ssrn.com/abstract=1636706

Nieuwenhuis, J.H. (2000), 'De Constitutie van het burgerlijk recht', *RM Themis* 2000, p. 203-211.

Annex I: Selected provisions about research exceptions to copyright

France
No implementing provision found[10]

Germany[11]
Implementing provisions only relate to exceptions for educational purposes
(viz. art 46, 47 and 52 German Copyright Act)

Italy
Art. 64 sexies
1. The authorization by the rightholder provided for in art. 64-quinquies shall not be required in the following cases:
a) where the database is accessed and visualized for the sole purpose of illustration for teaching or scientific research, as long as the source is indicated and to the extent justified by the non-commercial purposes to be achieved. Within the above activities of access and visualization, the possible operations of permanent reproduction of the contents in whole or in a large part on any other carrier shall always be subject to the rightholder's authorization:

Netherlands
Art. 12.5 Dutch Copyright Act

[10] Source: http://www.wipo.int/wipolex/en/text.jsp?file_id=180336 (visited 6 October 2011)

[11] Source: http://www.iuscomp.org/gla/statutes/UrhG.htm#46 (visited 6 october 2011)

A recitation, performance or presentation which is exclusively for the purposes of education provided on behalf of the public authorities or a non-profit-making legal person, in so far as such a recitation, performance or presentation forms part of the school work plan or curriculum where applicable, or which exclusively serves a scientific purpose, shall not be deemed public.[12]

Spain

No implementing provision found[13]

UK

Section 29 Copyright, Designs and Patents Act 1988

(1) Fair dealing with a literary, dramatic, musical or artistic work for the purposes of research for a non-commercial purpose does not infringe any copyright in the work provided that it is accompanied by a sufficient acknowledgement.

(1B) No acknowledgement is required in connection with fair dealing for the purposes mentioned in subsection (1) where this would be impossible for reasons of practicality or otherwise.

[12] Translation of Dutch Copyright Act by the Institute for Information Law, University of Amsterdam: http://www.ivir.nl/legislation/nl/copyrightact.html (visited 6 October 2011)

[13] Source: http://www.wipo.int/wipolex/en/text.jsp?file_id=126674, (visited 7 October 2011)